Towards Ethical Item Ranking: A Paradigm Shift from User-Centric to Item-Centric Approaches

Guilherme Ramos
Instituto de Telecomunicações, and
Instituto Superior Técnico, ULisboa
1049-001 Lisboa, Portugal
guilherme.ramos@tecnico.ulisboa.pt

Ludovico Boratto
University of Cagliari
Cagliari, Italy
ludovico.boratto@acm.org

Mirko Marras
University of Cagliari
Cagliari, Italy
mirko.marras@acm.org

ABSTRACT

Ranking systems are instrumental in shaping user experiences by determining the relevance and order of presented items. However, current approaches, particularly those revolving around user-centric reputation scoring, raise ethical concerns associated with scoring individuals. To counter such issues, in this paper, we introduce a novel item ranking system approach that strategically transitions its emphasis from scoring users to calculating item rankings relying exclusively on items' ratings information, to achieve the same objective. Experiments on three datasets show that our approach achieves higher effectiveness and efficiency than state-of-the-art baselines. Furthermore, the resulting rankings are more robust to spam and resistant to bribery, contributing to a novel and ethically sound direction for item ranking systems.

CCS CONCEPTS

• **Information systems** → **Learning to rank**; • **Applied computing** → *Law, social and behavioral sciences.*

KEYWORDS

Ranking System, Effectiveness, Efficiency, Bias, Non-discrimination, Robustness, Bribing.

ACM Reference Format:

Guilherme Ramos, Ludovico Boratto, and Mirko Marras. 2024. Towards Ethical Item Ranking: A Paradigm Shift from User-Centric to Item-Centric Approaches. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '24), July 14–18, 2024, Washington, DC, USA.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3626772.3657977

1 INTRODUCTION

The evolution of contemporary society into an information-driven economy has catalyzed a substantial surge in the expansion of e-commerce [12]. The swift and widespread dissemination of information on a global scale has ultimately redefined our social dynamics. Social networks and online forums, beyond fostering the exchange of opinions, have given rise to the influential phenomenon of online word of mouth (WOM) [15]. This dynamic shows considerable potential for propelling e-commerce sales [1, 4, 11, 14].

In this marketplace landscape, the effect of ratings and rankings has reached a crucial point for sellers. The visibility and sales performance of products/services are tied to the ratings they gather. Recent studies stress that, in numerous cases, online ratings (and corresponding rankings) have a more substantial influence than traditional marketing strategies [2, 6]. This influence leads to an increase in the adoption of rating manipulation strategies. Hence, we need to develop robust ranking systems (RS) resistant to spam [9].

Consequently, companies are increasingly investing resources in encouraging users to endorse their products or services [3]. This strategy may involve distributing product samples for users to provide authentic feedback or directly compensating users for generating positive ratings of their offerings while expressing negative feedback about those of competitors [7, 22]. The dynamic interplay between sellers and consumers in this landscape presents challenges, requiring the development of appropriate countermeasures. A notable solution to this issue is represented by RSs that leverage users' reputation to weight the items' ratings [5, 13, 16, 19, 24, 25].

While the aforementioned initiatives have positively impacted spam mitigation and addressed bribery concerns, they have concurrently sparked ethical considerations regarding the attribution of reputation to users. The underlying mechanisms may inadvertently contribute to biased rankings based on users' sensitive attributes and, therefore, be the source of discrimination [18, 20, 21] and misuse of personal data [26]. Moreover, the EU AI Act, as shown in Article 5(1)(c), forbids the deployment of intelligent systems for social scoring, thereby prohibiting the categorization or assessment of individuals based on social behavior or predicted attributes [17]. Thus, even existing bias mitigation procedures for RSs in [18, 20, 21] do not satisfy the EU AI Act since the approaches still assign a reputation to users. Initially, this rule is confined to actions made by public authorities, but amendments may extend it to private entities.

In this paper, we propose a novel ranking system grounded in a user-agnostic approach, thereby limiting the aforementioned pitfalls and establishing a more ethical system. The key intuition is to filter out and remove item's ratings that deviate from the majority. Through a comprehensive analysis on three different datasets,

SIGIR '24, July 14–18, 2024, Washington, DC, USA

ACM ISBN 979-8-4007-0431-4/24/07.
https://doi.org/10.1145/3626772.3657977

we highlight the merits of our user-agnostic approach in terms of effectiveness, efficiency, bribery resistance, robustness, and non-discrimination against state-of-the-art baselines. This paradigm shift holds the potential to cultivate more ethical online communities, conferring benefits across online platforms.

2 PROBLEM STATEMENT

Preliminaries. We denote a set of users by $\mathcal{U}$, a set of items by $\mathcal{I}$, and a set of permissible rating scores a user may assign to items by $\mathcal{R}$. We denote by $R \subset \mathcal{U} \times \mathcal{I} \times \mathcal{R}$ a set of ratings given by users to items, where $(u, i, r) \in R$ whenever user $u \in \mathcal{U}$ rated item $i \in \mathcal{I}$ with rating $r \in \mathcal{R}$, and if $(u, i, r) \in R$ then there is no $r' \in \mathcal{R}$ with $r' \neq r$ such that $(u, i, r') \in R$. In other words, the user could not assign more than one rating score to the same item. Also, for $i \in \mathcal{I}$, we denote the ratings given to item i by the entire user base by $R_i = \{r : (u, j, r) \in R \wedge j = i\}$ and the set of users that rated item i by $\mathcal{U}_i = \{u : (u, j, r) \in R \wedge j = i\}$. Without loss of generality, we assume that $\mathcal{R} \subset]0, 1]$. When this is not the case, we normalize each rating r to be $r' = (r - r_\perp + 1) / (r_\top - r_\perp + 1)$, where $r_\perp$ and $r_\top$ are, respectively, the smallest and largest ratings in the dataset[1]. In the non-personalized scenario, an RS aggregates ratings R_i given to a certain item i to determine the relevance score for that item i (e.g., based on the reputation of users rating it). The relevance scores are used to sort items using a certain criterion (e.g., descending order).

Ranking system goals. In the evolving regulatory landscape of RSs, there arises a critical need to explore alternative paradigms beyond (user-centric) reputation-based approaches [13, 20, 21]. Nonetheless, any replacement should carefully uphold the essential and desirable properties associated with the original reputation-based RSs. Therefore, given a dataset $R \subset \mathcal{U} \times \mathcal{I} \times \mathcal{R}$ containing ratings assigned by users to items, we aim to devise rankings for items in alignment with the following ranking system's properties:

- **RSP$_1$** *Efficiency*: item rankings computation should be reasonably rapid and adaptable to distributed processing [13, 24].
- **RSP$_2$** *Effectiveness*: item rankings should align with the ratings' arithmetic average, as per the original definition [13, 24].
- **RSP$_3$** *Spam Robustness*: item rankings should be robust to users attempting to manipulate the ratings via spamming [13, 24].
- **RSP$_4$** *Bribery Resistance*: item rankings should counteract bribery (i.e., seller incentives to inflate their ratings) [7, 24].
- **RSP$_5$** *Non-discrimination*: the items' rankings should not exhibit consistent discrimination on ratings of user(s) [18, 20, 21].[2]

3 THE PROPOSED RANKING SYSTEM

Next, we propose a user-agnostic RSs designed to intuitively filter out outlier ratings, without scoring users according to their estimated reputation. The stages of our approach are delineated in Algorithm 1. The foundational principle revolves around the systematic removal of ratings that substantially deviate from the average opinion about the item. This removal is crucial for an RS to be robust to spam and bribery [10, 13, 24, 25]. For each item, our approach computes the average and standard deviation for the ratings given by users. Then, ratings away from the average are discarded. The process repeats until reaching a fixed point. Note that the generated item rankings are independent of the users providing ratings for items (i.e., non-discriminatory), since no reputation score is assigned to each user; thus, we fulfill **RSP$_5$** by design.

Algorithm 1 User-agnostic ranking system (UARS)

1: **input**: set of ratings given by users to items, $R \subset \mathcal{U} \times \mathcal{I} \times \mathcal{R}$
2: **output**: set of items' rankings/scores $S \subset]0, 1]^{|\mathcal{I}|}$
3: **for** $i = 1, \ldots, |\mathcal{I}|$ **do**
4: **set** $R_i = \{r : (u, j, r) \in R \wedge j = i\}$, $ok \leftarrow False$
5: **while** $\neg ok$ **do**
6: **set** $\mu_i = \sum_{r \in R_i} r / |R_i|$ ▷ compute average and standard deviation
7: **set** $\sigma_i \leftarrow 0$ if $(|R_i| == 1)$ else $\sqrt{\frac{1}{|R_i|-1} \sum_{r \in R_i} (r - \mu_i)^2}$
8: **set** $R'_i \leftarrow \{r \in R_i : (r - \mu_i)^2 \leq \sigma_i\}$ ▷ discard ratings away from μ_i
9: **set** $ok \leftarrow (|R_i| == |R'_i|)$ ▷ check for a fixed point
10: **set** $R_i \leftarrow R'_i$ ▷ update the set of considered ratings
11: **end while**
12: **set** $s_i \leftarrow \sum_{r \in R_i} r / |R_i|$ ▷ set the ranking as the average of the filtered ratings
13: **end for**
14: **set** $S = \{s_i : i \in \mathcal{I}\}$

To motivate the idea behind our approach, consider an item i with users' ratings of $\{3, 4, 5, 5, 4, 1, 2, 5, 5, 4\}$. The re-scaled ratings are $R_i = \{0.6, 0.8, 1, 1, 0.8, 0.2, 0.4, 1, 1, 0.8\}$. Thus, $\mu_i = 0.76$ and $\sigma_i \approx 0.2797$, and, therefore, after the first iteration of the while loop (steps 7–11), $R_i = \{0.6, 0.8, 1, 1, 0.8, 0.4, 1, 1, 0.8\}$ (the smallest rating is eliminated). In the next iteration R_i does not change and the final ranking of item i becomes $s_i \approx 0.82$. Next, we prove that Algorithm 1 converges and detail its computational complexity.

Theorem 3.1. *Algorithm 1 converges.* ∘

Proof. The proof easily follows from noticing that, for each item, the set of ratings is filtered in the while loop of Algorithm 1. Therefore, the size of the set cannot increase. Thus, it can remain the same (and the while loop terminates) or it can decrease. However, the minimum size that the set can assume is 1, since in such case all the values in the set have distance zero to the average which is less or equal to the standard deviation (that is zero). □

Proposition 1. *The computational complexity of Algorithm 1 is $O\left(|\mathcal{I}||R_\alpha|^2\right)$, where $\alpha = \arg\max_{i \in \mathcal{I}} |R_i|$.* ∘

Proof. The main `for` loop, steps 3–13, runs $|\mathcal{I}|$ times. Then, the `while` loop, steps 5–11, processes the elements in R_i until reaching a fixed point, i.e., until R_i remain unchanged. At most, the `while` loop can have $|R_i|$ iterations. Additionally, steps 7, 8 and 9 iterates through all the elements in R_i. Thus, steps 5–11 have worst-case computational complexity of $O\left(|R_\alpha|^2\right)$, where $\alpha = \arg\max_{i \in \mathcal{I}} |R_i|$. Therefore, the total computational complexity is $O\left(|\mathcal{I}||\mathcal{R}_\alpha|^2\right)$. □

Note that Algorithm 1 is well-suited to a distributed or parallel setup due to the independence of instructions in the main `for` loop.

[1] We denote the number of elements in a set $\mathcal{S}$ by $|\mathcal{S}|$, the closed interval of real numbers between $a \in \mathbb{R}$ and $b \in \mathbb{R}$, with $a \leq b$ by $[a, b] \equiv \{x : x \in \mathbb{R} \wedge a \leq x \leq b\}$, and the open intervals as $]a, b] \equiv [a, b] \setminus \{a\}$ and $[a, b[\equiv [a, b] \setminus \{b\}$.
[2] This means that the scoring/ranking process should be fair and impartial, not favoring/penalizing users based on an algorithmically-determined reputation score.

4 EXPERIMENTAL RESULTS

Here, we conduct a thorough analysis of our RS (UARS), labeled as Ours, across three real-world datasets[3]. Our evaluation includes a comprehensive performance comparison with a state-of-the-art user reputation-based RS, denoted as UR (User Reputation)[4], as introduced in [20]. For each dataset, we compare our RS UARS (Ours) with UR regarding efficiency, effectiveness (comparing both with the arithmetic average of ratings (AA)), spam robustness, and bribing resistance. Notably, the baseline RS fails to ensure $\mathbf{RSP}_5$, whereas our RS inherently satisfies it by design. Following this, we provide a thorough comparison focusing on $\mathbf{RSP}_1$–$\mathbf{RSP}_4$, previously proven to be met by the baseline in prior works [24, 25]. This analysis aims to show the advantages of our RS UARS across these four goals.

Datasets. To ensure replicability and adherence to recognized protocols, we use the three publicly available datasets:

- MovieLens-100k (ML-100k), which has 100 000 ratings from $|\mathcal{U}| = 943$ users to $|\mathcal{I}| = 1682$ items [8];
- MovieLens-1M (ML-1M), which has 1 000 209 ratings from $|\mathcal{U}| = 6\,040$ users to $|\mathcal{I}| = 3\,952$ items [8];
- Amazon Musical Instruments, which has 500 176 ratings from $|\mathcal{U}| = 339\,231$ users to $|\mathcal{I}| = 83\,046$ items [23].

Comparison on efficiency. Table 1 provides a comparison on the computational time, in seconds, required to compute item rankings across the datasets. Notably, the reported times for our RS UARS do not include parallel computation. Throughout all experiments, UARS invariably outperforms UR, being at least five times faster. This advantage highlights our RS's merits in better meeting $\mathbf{RSP}_1$.

Comparison on effectiveness. To evaluate the effectiveness of the rankings, we compare the final ranking distributions between our RS UARS (Ours) and UR using the simple arithmetic average across the three datasets. Figures 1 (a)–(c) shows the ranking distributions. While the arithmetic average (AA) yields optimal rankings, it is widely acknowledged to be susceptible to spam or bribery [13, 24, 25]. Both our RS UARS (Ours) and UR generate rankings that closely mirror the simple arithmetic average, underscoring their effectiveness and addressing $\mathbf{RSP}_2$. Note that, across the datasets, UR has between 3% and 5% difference in ranking distribution regarding AA, and Ours has between 5% and 7% difference regarding AA. More importantly, despite these small losses in effectiveness, Ours is more robust to spam and bribery, as shown below.

Comparison on robustness. To investigate the robustness of the proposed RS, simulations were executed involving the introduction of spamming users into each dataset. The aim was to assess the system's performance when faced with users exhibiting spam-like behavior, wherein ratings are assigned arbitrarily without regard for inherent quality or the user's genuine preference.

To evaluate the robustness to spamming users, we use the *robustness* τ that is defined as the Kendall's τ between the rankings of the dataset without spam and the rankings of the dataset with added spam [13, 24] – the larger the value, the more robust the ranking (↗). That is $\tau\left(S, S_{\text{spam}}\right)$, where $S = \{s_i \ : \ i \in \mathcal{I}\}$ are the

[3] Code implementation: https://github.com/xuizy/Ethical_Item_Ranking/tree/main.

[4] This RS has been proven to beat all the others in prior works [24, 25].

	ML-100k	ML-1M	Amazon Musical Inst.
UR	1.23s	12.70s	13.00s
Ours	**0.01s**	**1.03s**	**2.56s**

Table 1: Average computational times in seconds.

item rankings computed with an RS with the original dataset and where $S_{\text{spam}} = \{s_i \ : \ i \in \mathcal{I}\}$ are the item rankings computed with the same RS with the inclusion of spamming users in the original dataset. The Kendall's τ is a measure of association for assessing the concordance between two rankings or orders of data, and

$$\tau\left(S, S_{\text{spam}}\right) = \frac{\text{number of concordant pairs} - \text{number of discordant pairs}}{\text{total number of pairs}},$$

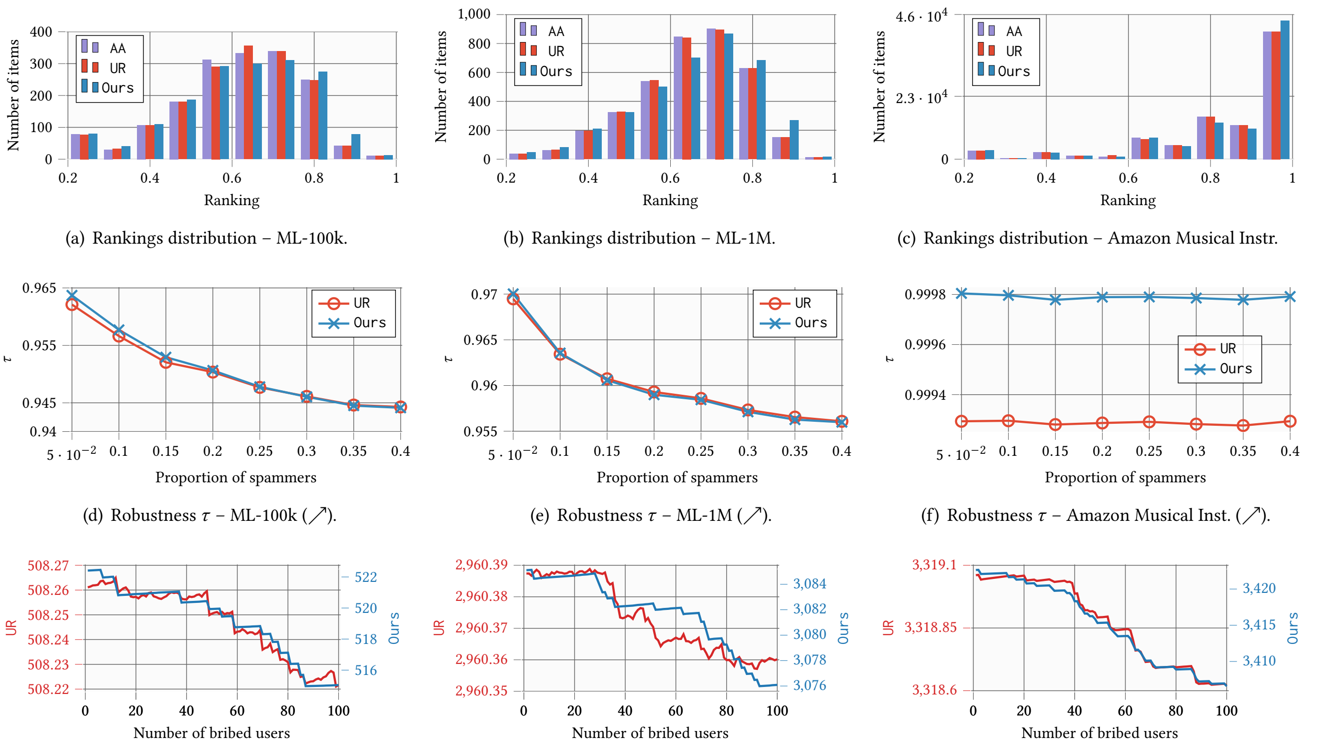
where *concordant pairs* are pairs of elements that maintain the same relative order in both rankings, while *discordant pairs* are those that have different relative orders. Hence, a τ equal to 1 means all comparisons are concordant, a τ equal to -1 means all comparisons are discordant, and a τ equal to 0 means no linear association.

In Figures 1 (d)–(f), we present Kendall's τ (the higher the more robust) values for both our RS UARS and the baseline UR across the three datasets. Thus, UARS demonstrates robustness comparable to the baseline UR. In some instances, it even displays higher robustness, thereby effectively fulfilling $\mathbf{RSP}_3$. This occurs for the Amazon dataset, the one with larger ration between $|\mathcal{U}|$ and $|\mathcal{I}|$.

Comparison on bribing resistance. Finally, we assess the resistance of our RS UARS to bribery in comparison to UR. To conduct this evaluation, we simulate the scenario where the seller of each dataset's most-rated item sequentially bribes the first 100 users who rated the item with a rating below the maximum allowed to change the rating to the maximum one. Then, we compute the profit after bribing each user (i.e., the wealth increase/decrease). The larger the wealth decreases the better the bribery resistance (Δ ↗).

To assess the profit, we use the definitions in [25]. We denote the *wealth* of item i seller by J_i, and it is computed as $J_i = |\mathcal{U}_i| s_i$. The seller's strategy for item i involves targeting a specific group of users and allocating its resources to encourage them to either provide ratings for or enhance their existing ratings of item i. An *elementary* bribing strategy ς_u^i is when item i seller targets a single user u to bribe to either rate item i with ρ_u or change the given rating by ρ_u. In all cases, $0 \leq \rho_u + r_{ui} \leq 1$, where r_{ui} is the rating user u gave to item i and $r_{ui} = 0$ if user u did not rate item i. A *compound* strategy is the composition of two or more elementary bribing strategies, i.e., $\varsigma^i_{u_1,\ldots,u_k} = \varsigma^i_{u_1} \circ \ldots \circ \varsigma^i_{u_k}$, where $\{u_1, \ldots, u_k\} \subset \mathcal{U}$. The wealth that seller of item i spends to perform the bribing strategy $\varsigma^i \equiv \varsigma^i_{u_1,\ldots,u_k}$ is $[\![\varsigma^i]\!] \equiv \sum_{j=1}^{k} |\rho_{u_j}|$. After performing such bribing strategy, wealth of item i seller becomes $J_{\varsigma^i} = |\mathcal{U}_{\varsigma^i}| s_{\varsigma^i} - [\![\varsigma^i]\!]$, where $\mathcal{U}_{\varsigma^i}$ is the set of users that rated item i after bribing strategy ς^i, and s_{ς^i} is the ranking of item i after bribing strategy ς^i. Therefore, the profit of such bribing strategy is calculated as $\pi_{\varsigma^i} = J_{\varsigma^i} - J_i$. The strategy is profitable if $\pi_{\varsigma^i} > 0$ and non-profitable otherwise.

Figures 1 (g)–(i) summarize the obtained wealth scores (the higher the decrease in wealth, the more resistant). Firstly, we observe that, for the majority of users, it is not financially advantageous to bribe them. Secondly, we note that the wealth decrease has a larger impact on our proposal. For ML-100k, the wealth decreases



(a) Rankings distribution – ML-100k. (b) Rankings distribution – ML-1M. (c) Rankings distribution – Amazon Musical Instr.

(d) Robustness τ – ML-100k (↗). (e) Robustness τ – ML-1M (↗). (f) Robustness τ – Amazon Musical Inst. (↗).

(g) Bribing wealth – ML-100k (Δ ↗). (h) Bribing wealth – ML-1M (Δ ↗). (i) Bribing wealth – Amazon Musical Instr. (Δ ↗).

Figure 1: Experimental results using UARS (Ours) and UR. AA denotes the arithmetic average of ratings. The first row presents results for effectiveness, the second one for robustness, and the third one for bribing resistance.

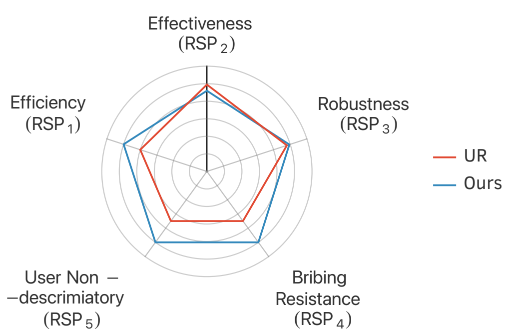


Figure 2: Summary comparison of our RS w/ the baseline, the comparison is relative, to depict which properties are improved with our proposal (Ours).

by approximately 7 units in UARS compared to a decrease of 0.5 in UR. For ML-1M, the wealth decreases by about 9 units in UARS versus a decrease of 0.3 in UR. Similarly, in Amazon Musical Instruments, the wealth decreases by around 20 units in UARS compared to a decrease of 0.4 in UR. So, we better achieved $\textbf{RSP}_4$.

Results summary. In Figure 2, we present a Kiviat chart that digests the results regarding our research goals. Thus, UARS (Ours) meets a better trade-off than the baseline on the selected goals.

5 CONCLUSIONS AND FUTURE WORK

In conclusion, our proposed RS emerges as a promising solution to address the ethical complexities inherent in user reputation scoring. The introduction of a user-agnostic ranking approach not only improves computational efficiency but also establishes a system capable of seamless computation in both distributed and parallel frameworks, preserving the efficacy of rankings. Moreover, our RS demonstrates more robustness against spamming users compared to the current state-of-the-art user reputation-based RS. Also, we illustrate that our proposal showcases superior resistance to bribery when compared to existing methods. Notably, our approach significantly contributes to the foundation of more trustworthy RSs.

In the next steps, a more comprehensive investigation into the profitability of specific bribing strategies is planned. Additionally, we will incorporate diverse real-world datasets and the integration of machine learning techniques for predictive modeling of user behaviors and preferences. Finally, considering the dynamic nature of online interactions, an investigation into the adaptability of the proposed RS to evolving user behaviors and emerging patterns would contribute to its robustness.

REFERENCES

[1] Barbara Bickart and Robert M Schindler. 2001. Internet forums as influential sources of consumer information. *Journal of interactive marketing* 15, 3 (2001), 31–40.

[2] Judith A Chevalier and Dina Mayzlin. 2006. The effect of word of mouth on sales: Online book reviews. *Journal of marketing research* 43, 3 (2006), 345–354.

[3] Robert B Cialdini. 2021. *Influence et manipulation: L'art de la persuasion.* First, Brussels, Belgium.

[4] Alanah Davis and Deepak Khazanchi. 2008. An empirical study of online word of mouth as a predictor for multi-product category e-commerce sales. *Electronic markets* 18, 2 (2008), 130–141.

[5] Cristobald De Kerchove and Paul Van Dooren. 2010. Iterative filtering in reputation systems. *SIAM J. Matrix Anal. Appl.* 31, 4 (2010), 1812–1834.

[6] Peter De Maeyer. 2012. Impact of online consumer reviews on sales and price strategies: A review and directions for future research. *Journal of Product & Brand Management* 21, 2 (2012), 132–139.

[7] Umberto Grandi and Paolo Turrini. 2016. A network-based rating system and its resistance to bribery. In *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence*. International Joint Conferences on Artificial Intelligence Organization, California, USA, 301–307.

[8] F Maxwell Harper and Joseph A Konstan. 2015. The movielens datasets: History and context. *Acm transactions on interactive intelligent systems (tiis)* 5, 4 (2015), 1–19.

[9] Nan Hu, Indranil Bose, Noi Sian Koh, and Ling Liu. 2012. Manipulation of online reviews: An analysis of ratings, readability, and sentiments. *Decision support systems* 52, 3 (2012), 674–684.

[10] Nan Hu, Paul A Pavlou, and Jennifer Zhang. 2006. Can online reviews reveal a product's true quality? Empirical findings and analytical modeling of online word-of-mouth communication. In *Proceedings of the 7th ACM conference on Electronic commerce*. Association for Computing Machinery, New York, NY, USA, 324–330.

[11] Jan Kietzmann and Ana Canhoto. 2013. Bittersweet! Understanding and managing electronic word of mouth. *Journal of Public Affairs* 13, 2 (2013), 146–159.

[12] Kenneth C Laudon and Carol Guercio Traver. 2020. *E-commerce 2019: Business, technology, society.* Pearson, United Kingdom.

[13] Rong-Hua Li, Jeffery Xu Yu, Xin Huang, and Hong Cheng. 2012. Robust reputation-based ranking on bipartite rating networks. In *Proceedings of the 2012 SIAM international conference on data mining*. SIAM, Society for Industrial and Applied Mathematics (SIAM), Anaheim, California, USA, 612–623.

[14] Ewa Maslowska, Edward C. Malthouse, and Stefan F. Bernritter. 2017. *The Effect of Online Customer Reviews' Characteristics on Sales.* Springer Fachmedien Wiesbaden, Wiesbaden, 87–100. https://doi.org/10.1007/978-3-658-15220-8_8

[15] Kelley O'Reilly and Sherry Marx. 2011. How young, technical consumers assess online WOM credibility. *Qualitative Market Research: An International Journal* 14, 4 (2011), 330–359.

[16] Larry Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. 1997. *PageRank: Bringing order to the web.* Technical Report. Stanford Digital Libraries Working Paper.

[17] European Parliament. 2023. Artificial Intelligence Act. https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236_EN.pdf. Accessed: 2024-02-06.

[18] Guilherme Ramos and Ludovico Boratto. 2020. Reputation (in) dependence in ranking systems: Demographics influence over output disparities. In *Proceedings of the 43rd international ACM SIGIR conference on Research and Development in Information Retrieval*. Association for Computing Machinery, New York, NY, USA, 2061–2064.

[19] Guilherme Ramos, Ludovico Boratto, and Carlos Caleiro. 2020. On the negative impact of social influence in recommender systems: A study of bribery in collaborative hybrid algorithms. *Information Processing & Management* 57, 2 (2020), 102058.

[20] Guilherme Ramos, Ludovico Boratto, and Mirko Marras. 2021. Reputation equity in ranking systems. In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*. Association for Computing Machinery, New York, NY, USA, 3378–3382.

[21] Guilherme Ramos, Ludovico Boratto, and Mirko Marras. 2022. Robust reputation independence in ranking systems for multiple sensitive attributes. *Machine Learning* 111, 10 (2022), 3769–3796.

[22] Guilherme Ramos, Daniel Silvestre, and Carlos Silvestre. 2021. Node and network resistance to bribery in multi-agent systems. *Systems & Control Letters* 147 (2021), 104842.

[23] Amazon Musical Instruments Reviews. 2020. Dataset for musical instruments review and ratings in Kaggle. https://www.kaggle.com/eswarchandt/amazon-music-reviews. Accessed: 2024-02-06.

[24] João Saúde, Guilherme Ramos, Ludovico Boratto, and Carlos Caleiro. 2021. A robust reputation-based group ranking system and its resistance to bribery. *ACM Transactions on Knowledge Discovery from Data (TKDD)* 16, 2 (2021), 1–35.

[25] João Saúde, Guilherme Ramos, Carlos Caleiro, and Soummya Kar. 2017. Reputation-based ranking systems and their resistance to bribery. In *2017 IEEE International Conference on Data Mining (ICDM)*. IEEE, Institute of Electrical and Electronics Engineers (IEEE), New Orleans, Luisiana, USA, 1063–1068.

[26] Sandra Wachter. 2018. Normative challenges of identification in the Internet of Things: Privacy, profiling, discrimination, and the GDPR. *Computer law & security review* 34, 3 (2018), 436–449.